1 Introduction

Deployable Aerial Communications systems have been used for many years by the armed services community for many years. There is reason to believe that some of the same strategies used by these agencies would prove valuable when applied to disaster recovery situations. The comments in this white paper are intended to suggest how to optimize those strategies for Disaster recovery efforts. The approach taken in this paper is to introduce a series of potential issues, and then to consider how this might affect how to approach the problem.

2 Issues

2.1 Multi-Purpose vs Dedicated Purpose

At the risk of beginning with a broad brush statement that can be proven to be incorrect in some small details, let's begin with the statement "every wireless communications system is designed and optimized for a specific use". That means the selection of a specific RF frequency of operation, bandwidth, modulation technique, media access control protocol, routing protocol, etc. Although we are deploying some Software Defined Radios, most deployed systems are still designed for a single purpose. Even with a Software Defined Radio technology, the agency that commissions the design of the radio makes all of the design choices, and there is always a limited set of resources to allocate to implement all of the available choices. What this means is that although there are many systems that might be useful in a disaster recovery scenario, many of them won't talk to other systems.

The problem could be as simple as choosing different RF frequencies and modulation schemes, or as insidious as having both chosen to implement an IP interface on their radio systems and then having chosen to use private IP addresses instead of public IP addresses. (Public IP addresses are scarce, they cost money to maintain, it is difficult to implement them in a mobile environment, etc.) Following this last thought one step further, you may ask why is the use of a private IP address significant. The reason is that it is forbidden to route traffic to and from private addresses across the public network (Network Address Translation is needed first). Where this line of logic is going is that there will need to be some kinds of gateways to allow the various systems to interconnect and build a large enough network to be effective.

So, the question here is who is to be responsible for defining what kinds of gateways are needed. There are no standards in this arena that can be referenced to claim that a particular kind of a gateway, or even one specific architecture for deploying gateways should be chosen over any other architecture. Some research in this area would be valuable. Ideally, the method of interconnecting each of these systems would be understood well in advance so that when the disaster occurs, it does not require a conference between the experts on each of those systems to derive a strategy ion the spot. Likewise, the process should not require manual changes to every device to make them compatible (say to change their IP addresses so that two devices aren't using the same address). This last problem is not hypothetical. Has anyone ever seen the address 192.168.0.1 on their home routers? This same kind of a behavior is not restricted to home users. Professional organizations do it too.

2.2 Mobile Networking is Different than Static Networking

When using Deployable Aerial Communications systems in a disaster recovery scenario, it seems only natural to want to run Internet services across it. Notwithstanding the huge success of the Internet, the

routing scheme used by the Internet does not gracefully manage mobility of its routers. In fact, the basic assumption of the Internet Protocol is that the IP address is both the identity of the host and its network location. That is to say, it is like a street address, where we navigate to the street by making a series of left and right turns at intersections until we reach the street, and finally the house number. When the street intersections keep changing, as they would for an airborne mobile network, the routing process reaches a point where it can no longer keep up with which street number is to its "left" or to its "right". In practical terms, no fully mobile airborne network (UAV speed) larger than 20 nodes has ever been built, and no land based (pedestrian speed) mobile network larger than 200 nodes has ever been built. That is not because people haven't tried to make it work. It is because the computation resources needed to track routes to individual users grows very quickly as the number of mobile routers increases. Designers of mobile networks where cost, size, weight and power are critical issues can't push the envelope too far.

Where this line of logic leads is that in a disaster recovery scenario, we cannot assume that simply interconnecting the routers from multiple smaller networks will be a solution to creating the infrastructure needed for thousands of users. We have to be more clever than that. In order to scale our architecture, we have to have allow the infrastructure to be composed of whatever sized mobile networks that each organization can bring to the table. Routing traffic between them and through them, while not forcing each of the smaller networks to have to join the same routing process, is our only hope for building an infrastructure of any size.

Before leaving this topic, it is worth noting that there may be some skeptical readers who observe that wireless cellular networks have grown to enormous sizes. This apparently disproves the assertion just made about deployed mobile network sizes being capped at only 20 or 200 nodes. Actually, it is precisely the difference between a static cell tower and a fully mobile UAV that causes the difference. In a cellular system cell towers don't move, only the user is moving. It is the motion of the cell towers that causes the incredible increase in resource usage to track individual users. Further, the mobile platforms are quite limited in the size, weight and power that those routers can apply to the problem, so they aren't able to keep up. We might look at this and just say that eventually the processing power will eventually increase to the point where larger sizes are practical. Alternatively, we could observe that most of the problem with tracking IP addresses in a mobile environment is self-inflicted (IP addressing was not designed to be mobile). There are clearly less resource intensive ways of solving this problem. We would do well to choose one of them and standardize on it as part of the DACA effort.

2.3 International Cooperation

Many disasters of are of such magnitude that help is offered by many nations of the world. Sometimes, those nations are willing to bring resources that could help build the communications infrastructure. (maybe even a deployable aerial communications systems). The interaction of nations is sometimes a little biased by fears of one nation (probably a friendly nation) spying on the systems of another nation (probably for economic gain). One of the constraints of any effective architecture is one that permits the resources of each nation to be used in a secure manner. Neither the host nation, nor any of the guest nations should be able to take advantage of the situation to access information they should not. They need to be faithful carriers of the data but they should not have full access to the networks of other nations.

2.4 Access to computation resources for Situational Awareness

Not only is it important for the architecture for a deployable aerial communications system to be able to support re-establishment of communications, it is also important that the system be able to provide

services. If we presume that the only purpose of building a network in a disaster situation is to connect back to the Internet, then we are presuming that we intend to carry all of the Internet traffic twice across the mobile network (once going and once returning). The capacity of the aerial network is certainly unlikely to be that of the wired network, so strategies that embed computing resources in the network itself would reduce the overall load on the network. Further, if those resources were assignable, then the location of the service could be deliberately selected to minimize traffic load.

3 Strategies for Success

Several of the strategies described below are part of a research program currently underway at Virginia Tech into the scalability of disaster recovery networks.

3.1 Backbone Networks

As suggested by the discussion of differences between mobile networking and static networking, building infrastructure to support many users with deployable aerial communication systems, i.e. building a mobile backbone, has to assume the use of multiple independent sub-networks. A logical follow-on assumption is that traversing the sub- network is accomplished by tunneling across it (probably encrypted) instead of translating every packet to the native addressing scheme of the sub-network.

With the assumption of tunneling across the backbone, we can admit the possibility of maintaining separation between the routing domains of each of the user groups that want to get services from the backbone. As illustrated in Figure 1, building the backbone requires two things. The first is that each sub-network that joins the backbone is able to provide a transparent bridging service to tunnel across the entire backbone. The second is that each sub-network that joins the backbone is able to support a transparent bridging service across its own segment of the backbone while maintaining a routing protocol between it and other segments of the backbone.

The implication of the first requirement is that there is a stable backbone wide addressing scheme that each sub-network of the backbone can apply to its own nodes. As discussed in section 2.2, the addressing scheme is not based on IP addresses. It could be, however, based on a node identifier assigned by the sub-network, and an identifier given to the sub-network by the overall network management function. In that case, routing across the backbone would only require each sub-network (Cluster) to know if the destination was one of its own nodes (i.e. the destination address has the same cluster identifier), and if so, then how to forward traffic to that node. It can be noted that with this strategy the individual sub-network does not need to track the location of individual nodes within another sub-network. It just needs to maintain a routing protocol with other sub-networks (Clusters) to track the location of other clusters. If we assume that the clusters are randomly moving with respect to each other, that suggests that we should be able to build a backbone with up to 20 segments. If we assume that we can assign the segments somewhat geographically, then the relationship of clusters to each other is stable, and we should be able to increase the number of segments (sub-networks) to at least 200.

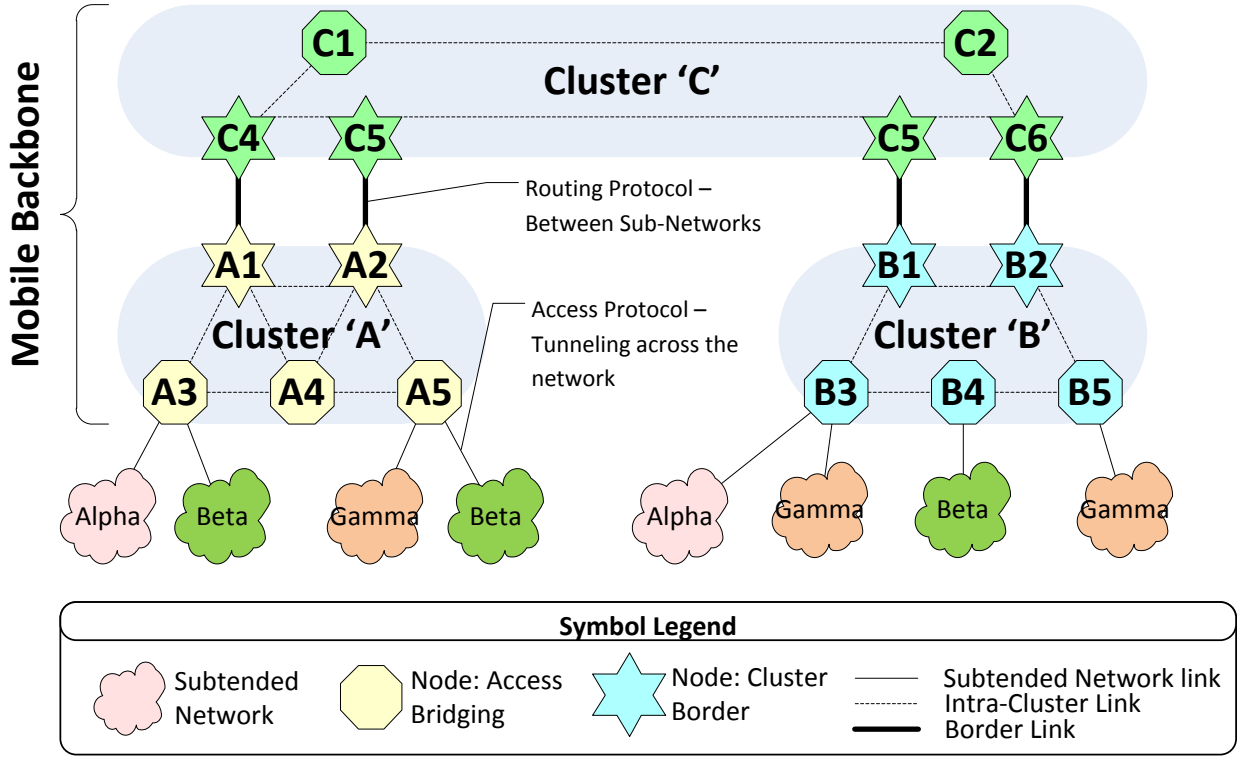

Figure 1 A Backbone is formed from multiple sub-networks (Clusters) and provides services to other wireless networks

To summarize, we can build a backbone network from smaller sub-networks. We can apply a backbone wide addressing scheme to nodes on the backbone. We can use that addressing scheme to develop a routing protocol that does not require each sub-network to track the location of individual nodes. That routing protocol does not succumb to the scaling problems of IP based routing protocols. We can use tunneling to traverse the network to maintain isolation between private and public networks. Standards body work needed to implement DACA includes the interface between segments of the backbone, and between the backbone and users of the backbone. The final question is how we efficiently deal with the question of letting other nodes know about where each mobile wireless node is located.

A partial answer to this question is that we exploit the same kinds of distinctions that are used to separate Verizon cellular customers from Sprint cellular customers, and Wi-Fi customers from Wi-Max customers, or P25 users from cell-phone users. Customers on those networks don't actually use each other's infrastructure, they use their own infrastructure and meet at the Internet. The mobile wireless backbone is not actually the Internet, it is more like the Virtual Local Area Networks that are used by a corporation to allow their workers to move to new cubicles without having re-wire their buildings. A logical strategy is to provide an access point to the actual Internet for each independent set of users. So, in Figure 1, one of the alpha network's connection points to the backbone would be an Internet access point, likewise for beta and gamma.

A more complete answer would note that each of the different kinds of networks listed in the last paragraph has a different kind of service architecture. Some of those types of networks require servers that would need to be embedded in the backbone, like they are embedded in the cell-tower networks today. That leads to the need for system flexibility that will be described in the next section.

3.2 Flexibly Reconfigurable Radios (Or is it Subsystems?)

Software Defined Radio has made great strides in being able to allow a single set of hardware to emulate the behavior of many different kinds of radios. Unless we want to build a separate backbone for Sprint customers, and for Verizon customers, and for AT&T customers, and for P25 customers, and for every other kind of customer, (like we have today in the wired world), the aerial nodes that we use need to be able to be configured to provide multiple services from the same platform. Conceptually, this is illustrated in Figure 2. The Access Bridging node of a backbone network has a routing function to be able to traverse the backbone network. It also has to have the individual access bridging functions needed for each type of user (e.g. Sprint, Verizon, Wi-Fi, Wi-Max, etc.). As already noted, that may mean that it has to accommodate a number of servers as well as radio elements.

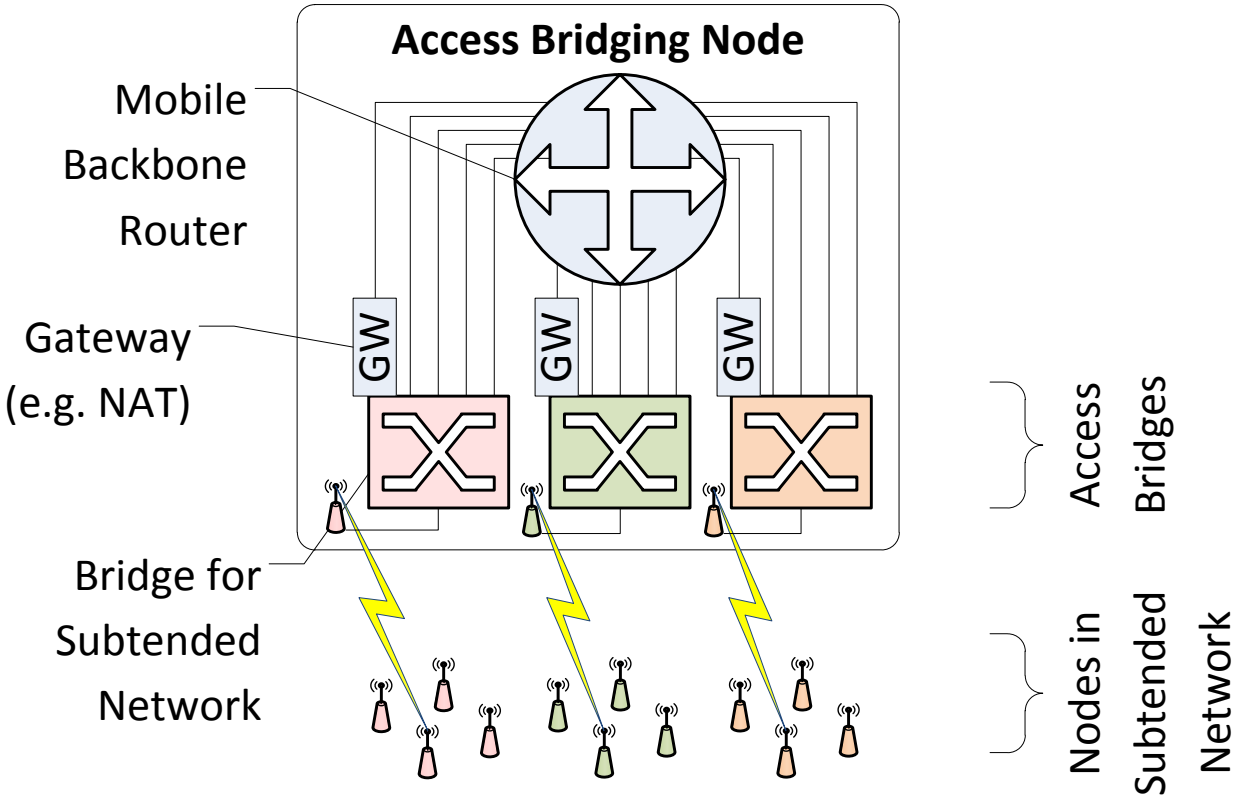

Figure 2 Flexibly Reconfigurable Nodes provide access to the backbone

Achieving this level of flexibility is something that is only now being demanded of radio systems, and is an emerging research area. The question of how a node provides a variable number of access bridging instances, with a variable number of servers and services, of types that need to be selected at deployment time (as opposed to system design time) is a problem that has not yet been fully solved. However, it is clear that operating an efficient disaster recovery backbone would require this behavior since the cost of fielding multiple systems would make it difficult to justify.

As a side benefit, the processing resources needed to provide a variable number of servers could also be put to other uses. For example, they could be used to form a cloud computing environment that would permit the processing of sensor data (maybe an air-dropped sensor network could be one of the users of the backbone). This mode of operation suggests that processing resources for a backbone node should not be permanently assigned to individual access bridge instances. Rather, they should be switched in on an as-needed basis so that they could be re-assigned as needed.

4 Summary and Conclusions

Building a disaster recovery network with aerial platforms is, in some ways, a new mode of operation for which current systems have not yet been designed. Most communications equipment is (has been) designed with a single type of network in mind and so is not able to flexibly accommodate the range of radio networks needed. To be able to leverage these assets in a disaster recovery scenario will require some standardization work to permit a backbone to be quickly constructed from independent systems, and some development work to select appropriate protocols. A conscious decision to fund this kind of research may be needed since the current owners of networks like this derive little benefit from supporting additional protocols if nobody else supports them.

To be able to operate an efficient backbone, there needs to be additional development work to manage resources at individual nodes. The design target for disaster recovery systems should be a fully reconfigurable node that is both capable of emulating various radio systems (i.e. has Software Defined Radio capabilities), and managing server assets. Although there some benefit to owners of current systems from the perspective of managing processor resources, the current purchasers of radios, and the purchasers of server gear are typically different organizations that don't coordinate their efforts. Inclusion of the requirement to dynamically reconfigure both radio and server resources in future contracts for aerial platforms would be a spur to help this process along.